

*B* structurally analyzing the [document] text, the method being implemented by a processor coupled to memory, the instructions comprising the steps of:

*(b)* 1) generating a cue vector from the text, the cue vector representing occurrences in the text of a first set of nonstructural, surface cues, the first set of cues including a punctuational cue;

2) determining a relevancy to the text of each facet of a second set of facets using the cue vector and a weighting vector; and

3) identifying from a third set of text genre types a text genre type of the text based upon those facets of the second set that are relevant to the text.

---

REMARKS

Claim 1-27 are pending. By this Amendment, claims 1, 5, 19, 22 and 27 are amended. For at least the following reasons, Applicants respectfully submit that all of claims 1-27 are allowable. Reconsideration of claims 1-27 in view of the foregoing amendments and the following remarks is respectfully requested.

I. FORMAL MATTERS

A. Item 1 of the Office Action objected to claim 22 due to a minor informality. Claim 22 has been amended to overcome this objection.

B. Item 3 of the Office Action rejects claim 5 under 35 U.S.C. §112, second paragraph, as being indefinite. Claim 5 has been amended to overcome this rejection. Specifically, claim 5 has been amended to depend on claim 4 instead of claim 1 to overcome the lack of antecedent basis. Reconsideration in view of this amendment is respectfully requested.

II. CLAIMS 1-27 DEFINE PATENTABLE SUBJECT MATTER

A. Item 5 of the Office Action rejects claims 1-3, 6, 13-14 and 26 under 35 U.S.C. §103(a) to U.S. Patent No. 5,913,185 to Martino et al. This rejection is respectfully traversed.

The Office Action asserts that Martino renders obvious all of the features recited in claims 1-3, 6, 13-14 and 26. However, Applicants respectfully submit that Martino does not teach or suggest a processor implemented method of identifying a document type of a document in machine readable form without structurally analyzing the document text, as recited in claim 1 and similarly in claims 19, 26 and 27.

That is, in claims 1, 19, 26 and 27 a processor identifies a document type without structurally analyzing the text. Applicants respectfully submit that Martino does not teach or suggest this feature.

Rather, Martino relates to a method of identifying a language shift point in a document that is written in a plurality of natural languages. Martino uses an interval that is moved through a text document in a computer memory. As the interval is moved through the document, the system calculates a probability that the text in the interval is written in each of a plurality of languages for that interval. Using this calculation, the unit assigns a first and second candidate language. The first candidate language is identified if it has the highest probability of all the candidate languages within that interval. Once the two candidate languages have been assigned, a language shift point in the document is identified where the relative probability of a second candidate language is higher than the current language at a new position of the interval. At that point, the second candidate language is classified as a current language in the document after the language shift point. The process continues to identify other language shift points in the document.

Martino uses the shift point to determine approximate location of language change within the document. To more precisely calculate where these shifts occur, Martino uses a log file of the scanned text which is examined for punctuation, capitalization, font changes or paragraph changes. See Martino, column 11, lines 8-11. This process is repeated for each approximated shift point. Applicants respectfully assert that Martino does not disclose the various features outlined above, as recited in claims 1, 19, 26 and 27.

In view of the foregoing remarks, Applicants respectfully submit that Martino does not teach or disclosure the subject matter recited in claims 1-27. Thus, Applicants respectfully submit that claims 1-3, 6, 13-14 and 26 define subject matter patentable over the prior art. Withdrawal of the rejections of claims 1-3, 6, 13-14 and 26 under 35 U.S.C. §103(a) is respectfully solicited.

B. Item 6 of the Office Action rejects claims 4, 5, 7-12, 15-25 and 27 under 35 U.S.C. §103(a) as obvious to Martino and further in view of "The Multi-Dimensional Approach to Linguistic Analysis of Genre Variation: An Overview of Methodology and Findings", by Biber. This rejection is respectfully traversed.

Applicants respectfully submit that Biber does not make up for the deficiencies of Martino outlined above with respect to claim 1-27. Biber discloses a summary of the major methods and results of multi-dimensional approach to genre variation. This approach includes a combination of resources of computational tools, large text corpora, and multivariate statistical tubes. Applicants assert that Biber does not disclose a process or implemented method of identifying a document type of a document in machine readable form within structurally analyzing the document text.

As neither Martino nor Biber teaches, discloses or suggests a processor implemented method of identifying a document type of a document in machine readable form without structurally analyzing the document text, a combination of Martino and Biber fails to teach,

disclosure or suggest every feature of claims 1-27. Thus the combination of Martino and Biber fails to render obvious the subject matter of claims 1-27. Withdrawal of the rejection of claims 4, 5, 7-12, 15-25 and 27 under 35 U.S.C. §103(a) as obvious in view of the combination of Martino and Biber is respectfully solicited.

II. Conclusion

In view of the foregoing remarks, Applicants respectfully submit that this application is in condition for allowance. Favorable reconsideration and prompt allowance are earnestly solicited. Should the Examiner believe that anything further is desirable in order to place this application in better condition for allowance, the Examiner is requested to contact the Applicants undersigned representative at the telephone number listed below.

Respectfully submitted,



James A. Oliff  
Registration No. 27,075

Thomas J. Pardini  
Registration No. 30,411

JAO:CCH/lrt

Date: August 14, 2000

**OLIFF & BERRIDGE, PLC**  
**P.O. Box 19928**  
**Alexandria, Virginia 22320**  
**Telephone: (703) 836-6400**

<p>DEPOSIT ACCOUNT USE AUTHORIZATION Please grant any extension necessary for entry; Charge any fees due to Xerox Deposit Account No. 24-0037</p>
---